Genetic diversity and population structure analyses in the Alpine plum (Prunus brigantina Vill.) confirm its affiliation to the Armeniaca section

Liu Shuo, Decroocq Stephane, Harte Elodie, Tricon David, Chague Aurelie, Balakishiyeva Gulnara, Kostritsyna Tatiana, Turdiev Timur, Fisher-Le Saux Marion, Sylvie Dallot, et al.

1 **Genetic diversity and population structure analyses in the Alpine plum**

2 **(*Prunus brigantina* Vill.) confirm its affiliation to the Armeniaca section**

3 Liu Shuo[1,2], Decroocq Stephane[1], Harte Elodie[1], Tricon David[1], Chague Aurelie[1],

4 Balakishiyeva Gulnara[3], Kostritsyna Tatiana[4], Turdiev Timur[5], Fisher-Le Saux Marion[6],

5 Dallot Sylvie[7], Giraud Tatiana[8], Decroocq Veronique [1,Ŧ]

6

7 [1] Univ. Bordeaux, INRAE, Biologie du Fruit et Pathologie, UMR 1332, F-33140, Villenave

8 d'Ornon, France.

9 [2] Liaoning Institute of Pomology, Tiedong Street, Xiongyue Town, Bayuquan District,

10 Yingkou City, Liaoning, 115009, China

11 [3] Institute of Molecular Biology and Biotechnology, Azerbaijan National Academy of

12 Sciences, 11 Izzat Nabiev str., AZ-1073, Baku, Azerbaijan

13 [4] International Higher School of Medicine, 1F Intergelpo Street, 720054 Bishkek, The

14 Kyrgyzstan

15 [5] Institute of Plant Biology and Biotechnology, Timiryazev Str. 45, 050040 Almaty,

16 Kazakhstan

17 [6] IRHS, AgroCampus-Ouest, INRAE, Université d'Angers, SFR 4207 QuaSav, 49071,

18 Beaucouzé, France

19 [7] BGPI, Univ Montpellier, INRAE, CIRAD, Institut Agro, F-34398 Montpellier cedex 5,

20 France

21 [8] Ecologie Systematique Evolution, CNRS, AgroParisTech, Université Paris-Saclay,

22 91400 Orsay, France

23

24 [Ŧ] Corresponding authors: veronique.decroocq@inrae.fr

25

26 **Abstract**

27 In-depth characterization of the genetic diversity and population structure of wild relatives

28 of crops is of paramount importance for genetic improvement and biodiversity

29 conservation, and is particularly crucial when the wild relatives of crops are endangered.

30 In this study, we therefore sampled the Alpine plum (Briançon apricot) *Prunus brigantina*

31 Vill. across its natural distribution in the French Alps, where its populations are severely

32 fragmented and its population size strongly impacted by humans. We analysed 71 wild *P.*

33 *brigantina* samples with 34 nuclear markers and studied their genetic diversity and

34 population structure, with the aim to inform *in situ* conservation measures and build a core

35 collection for long-term ex-situ conservation. We also examined the genetic relationships

36 of *P. brigantina* with other species in the Prunophora subgenus, encompassing the Prunus

37 (Eurasian plums), Prunocerasus (North-American plums) and Armeniaca (apricots)

38 sections, to check its current taxonomy. We detected a moderate genetic diversity in *P.*

39 *brigantina* and a Bayesian model-based clustering approach revealed the existence of

40 three genetically differentiated clusters, endemic to three geographical regions in the Alps,

41 which will be important for *in situ* conservation measures. Based on genetic diversity and

42 population structure analyses, a subset of 36 accessions were selected for *ex-situ*

43 conservation in a core collection that encompasses the whole detected *P. brigantina* allelic

44 diversity. Using a dataset of cultivated apricots and wild cherry plums (*P. cerasifera*)

45 genotyped with the same markers, we detected gene flow neither with European *P.*

46 *armeniaca* cultivars nor with diploid plums. In contrast with previous studies, dendrograms

47 and networks placed *P. brigantina* closer to Armeniaca species than to Prunus species.

48 Our results thus confirm the classification of *P. brigantina* within the Armeniaca section; it

49 also illustrates the importance of the sampling size and design in phylogenetic studies.

50

51 **Keywords**: Apricot, Prunus, classification, genetic structure, core collection, taxonomy

52

Introduction

Many wild crop relatives are endangered, because of fragmented and reduced habitats as well as crop-to-wild gene flow (Cornille et al. 2013b). In order to protect the biodiversity of wild crop relatives, we need to understand their population subdivision and genetic diversity distribution (Allendorf et al. 2012; Fahrig 2003). Studying the genetic diversity of crop-related species is not only important for biodiversity conservation but also for the sustainable use of valuable genetic resources through the set-up of *ex-situ* germplasm collections (Li and Pritchard 2009). Developing such collections requires obtaining a sufficient number of individuals to be representative of the species diversity (Frankel and Brown 1983; Glaszmann et al. 2010; Govindaraj et al. 2014). Core collections of woody perennial species have the additional advantages of being propagated vegetatively and maintained for decades, as clonemates, in field collections (Escribano et al. 2008).

Within the genus *Prunus* L. (stone fruit species), the subgenus Prunus (also called Prunophora Neck. to avoid confusion with the Prunus section and the *Prunus* genus) includes three sections: the Eurasian and North American plums (sections Prunus and Prunocerasus Koehne, respectively) and apricots (section Armeniaca (Mill.) K. Koch), which are all native from the Northern hemisphere (Rehder 1940) (Figure 1). It was shown that an ancient radiation of the *Prunus* genus through the Old and New Worlds and independent dispersal events across the North-American and Eurasian continents gave rise to, on one side, the Prunocerasus species, and on the other side, species of the Prunus and Armeniaca sections (Chin et al. 2014). Following the Rehder's classification, within the Prunophora subgenus the section Armeniaca comprises only diploid species, in which six species are recognized, based on morphological features: *P. armeniaca* L. (common apricot), *P. sibirica* L. (wild apricot in Northeastern Asia), *P. mandshurica* Maxim. (Northeast China and Eastern Russia), *P. mume* (Sieb.) Sieb. & Zucc. (South China and Japan), *P. holosericeae* Batal (South-West China) and *P. brigantina* Vill. (Figure 1). The first five species all originate from Asia, ranging from Central to North-East Asia, while *P. brigantina* (synonym *P. brigantiaca*, http://www.theplantlist.org) is native from Europe (Villars 1786). This species still grows in wild, patchy thickets in the Alps along the border between France and Italy in the Northern Mediterranean area, where it is considered either as an apricot or a plum species (Hagen et al. 2002; Pignatti 1982).

Prunus brigantina, alternatively called the Briançon apricot or the Alpine plum, was first reported in the French book <*Histoire des Plants de Dauphiné*> (Villars 1786). It grows in arid places in shrub and sparse thickets in the Alps, above 1,400 m altitude. Like other *Prunus* species, *P. brigantina* is hermaphrodite and is pollinated by insects, it flowers in

3

88　May and its fruits ripen from August to September (Noble et al. 2015; Tison and De
89　Foucault 2014). In natural stands, *P. brigantina* trees grow 2 to 5 meters high with non-
90　spiny branches and have heart leaves with double-serrated teeth (Figure 2a). The full-
91　fledged drupe from *P. brigantina* has a small size and appears glabrous with yellowish
92　fruit skin (Figure 2a). In the Alps, *P. brigantina* fruits are collected by locals to make jam
93　(Couplan 2009), and their seeds used to be processed for oil production instead of olive
94　or almond (Dupouy 1959). It is locally called 'Marmottier' or 'Afatoulier' and is recognized
95　as an endemic fruit tree in Europe. Its small and fragmented distribution suggests that it
96　may be threatened. However, there is currently insufficient information available to
97　evaluate the current genetic diversity of *P. brigantina* or its population subdivision, which
98　could contribute to determine the potential threats to this species or its conservation status
99　(IUCN Red List) (Branca and Donnini 2011), and to develop and inform conservation
100　programs.

101　　　　Previous phylogenetic studies questioned the Rehder's classification of *P.*
102　*brigantina* in the Armeniaca section (Hagen et al. 2002; Reales et al. 2010; Takeda et al.
103　1998; Zhebentyayeva et al. 2019). However, only one or two *P. brigantina* samples were
104　analysed, that had not been collected *in situ* but instead obtained from germplasm
105　repositories such as the Kew Royal Botanical Garden (UK), the Czech national genetic
106　resources of Lednice, the French Centre of genetic resources at INRAE-Montfavet or the
107　Japanese Chiyoda experimental station. Because of the ability of *P. brigantina* to be
108　propagated by grafting and its interfertility with species from both sections Prunus and
109　Armeniaca, the analyzed trees could be clonemates or hybrids, and their origin was
110　unknown. Moreover, sampling only one or two individuals per species is known to lower
111　the accuracy of phylogenetic analyses (Heled and Drummond 2010; Wiens and Servedio
112　1997). The genetic relationships between *P. brigantina* and other apricot species in the
113　section Armeniaca therefore remain unclear.

114　　　　To provide useful guidelines for *P. brigantina* conservation, a critical first task is first
115　assessing its genetic diversity distribution and population subdivision, in order to assess
116　whether local specificities need to be conserved. A second important aspect is to clarify
117　the taxonomic position of the different populations of *P. brigantina*. Indeed, differentiated
118　populations assigned to a given Latin species may actually represent different species
119　placed far apart in phylogeny, as found recently for *P. sibirica* (Liu et al. 2019). However,
120　there is currently no robust data on the genetic diversity or population structure of *P.*
121　*brigantina*. In the current study, we therefore conducted extensive sampling of *P.*
122　*brigantina* in its natural habitat and genotyped samples using 34 nuclear markers (Liu et
123　al. 2019). We assessed the genetic diversity of *P. brigantina* and its population structure,

124 as well as its relationship with Eurasian Armeniaca and Prunus species. We questioned

125 the affiliation of *P. brigantina* to either the Armeniaca or Prunus sections. For this purpose,

126 we performed a population genetic analysis using datasets with the main members of the

127 Armeniaca and Prunus sections and five outgroup species from the North-American

128 section Prunocerasus. Based on our molecular data, we identified a collection of unique

129 genotypes and selected the best subset for building a *P. brigantina* core collection,

130 maximizing allelic diversity, which will be very useful for further *P. brigantina*

131 characterization and for stone fruit crop improvement. In contrast to previous phylogenetic

132 studies, our population trees and networks further confirmed that *P. brigantina* is closer to

133 apricot species than to plum species.

134

Materials and Methods

135 **Materials and Methods**

136 ***In situ P. brigantina* sampling**

137 A total of 71 wild *P. brigantina* trees were collected in 2017 from three sampling sites, in

138 southeast France, across the Alps (Figure 2a and Table S1). Young leaves and mature

139 fruits from each tree were collected for DNA extraction and seedling growth, respectively.

140 At least one seedling from each sampled tree was kept for possible inclusion in a core

141 collection.

142 Representatives of other species of the Armeniaca section (*P. armeniaca*, *P. mume*, *P.*

143 *sibirica, P. mandshurica*), including two *P. brigantina* accessions maintained by the Centre

144 of genetic resources at INRA-Montfavet but with unknown origin, were previously

145 described and genotyped with the same set of molecular markers (Liu et al. 2019).

146

147 ***In situ* and *ex situ* sampling of representatives of Prunus and Prunocerasus species**

148 Part of the plum and plum-related material analysed in this study was kindly provided by

149 the North-American national repository (ARS-USDA, Davis, California, USA), the

150 Bourran's collection of Prunus (Prunus Genetic Resources Centre or Prunus GRC,

151 France) or was collected *in situ*, between 2008 and 2019, in Azerbaijan (Caucasia),

152 Kazakhstan and Kyrgyzstan (Central Asia) (Table S1b). One *P. cerasifera* sample (X29)

153 was collected *in situ* in South-West of France, along the Garonne river (Le Tourne-

154 Langoiran) and another one in the French Alps (FR_070) (Table S1b). In total, 82 diploid

155 samples were genotyped in the current study, genotypes for polyploids being difficult to

156 analyse. The diploid samples included representatives from *P. cerasifera* (*N*=66) (or

157 'cherry plum', alternatively called 'myrobolan' in Europe, *P. divaricata* Ledeb. in Caucasia

158 and *P. sogdiana* in Central Asia) and other species from the *Prunophora* subgenus: *P.*

159 *mexicana* (*N*=1), *P. munsoniana* (*N*=1, also called *Prunus rivularis*), *P. maritima* (*N*=1), *P.*

160 *americana* (*N*=1) and *P. subcordata* (*N*=1). *P. salicina* (Japanese plum) samples (*N*=10)

161 were composed of five cultivated accessions which included one plumcot, a hybrid

162 between *P. salicina* and *P. armeniaca* (called 'Rutland' in the ARS-USDA database,

163 P0489) and five wild *P. salicina* accessions, sampled in China (Table S1b). *Prunus*

164 *cerasifera* accessions used in this study originated from Europe (*N*=13), from Caucasia

165 and Russia (*N*=29) and from Central Asia (more precisely from Kazakhstan and

166 Kyrgyzstan, *N*=24) (see Table S1b for details).

167

168 **DNA extraction, microsatellite markers and polymerase chain reaction (PCR)**

169 **amplification**

170 Genomic DNA was extracted as described previously (Decroocq et al. 2016), either from

171 lyophilized leaves, bark or fresh flowers. We used 34 microsatellite markers distributed

172 across the eight *P. armeniaca* chromosomes and showing good amplification success as

173 well as substantial polymorphism within the different species of the section *Armeniaca*

174 (Liu et al. 2019). The same set of microsatellite markers were used to amplify PCR

175 fragments in species of the Armeniaca (*P. brigantina* incl.), Prunus (*P. cerasifera* incl., see

176 supplemental information) and Prunocerasus sections. Detailed information on these

177 microsatellite markers, including their repeat motifs, sequences, and amplification

178 conditions are available in (Liu et al. 2019). PCR amplification and fragment size

179 genotyping were performed on an ABI PRISM 3730 (Applied Biosystems) as described

180 previously (Decroocq et al. 2016). Alleles were scored with the GENEMAPPER 4.0

181 software (Applied Biosystems).

182

183 *Analyses of population subdivision and genetic relationships*

184 To assess the probability of observing unrelated individuals with the detected similar

185 genotypes given the population allelic frequencies, we used GENODIVE and the

186 corrected *Nei's* diversity estimate with a threshold of 50 (Meirmans and Van Tienderen

187 2004). We later retained only one individual of each pair detected as clonemates or

188 siblings for further analyses.

189 We identified population subdivision with the STRUCTURE software v. 2.3.3 (Pritchard et

190 al. 2000), without the use of *a priori* grouping information and assuming that individuals

6

had mixed ancestry with correlated allele frequencies among populations. The clustering method implemented in STRUCTURE is based on Monte Carlo Markov Chain (MCMC) simulations and is used to infer the proportion of ancestry of genotypes in K distinct clusters. We simulated K values ranging from 2 to 10 for the *P. brigantina* population and three additional datasets (Table 1 and S1c), obtained with the same genetic markers on Armeniaca and Prunus species originating from Central and Eastern Asia (Liu et al. 2019) (Supplemental information). For each K, we ran 10,000 generations of 'burn-in' and 100,000 MCMC. Simulations were repeated 10 times for each K value; the resulting matrices of estimated cluster membership coefficients (Q) were permuted with CLUMPP (Jakobsson and Rosenberg 2007). STRUCTURE barplots were displayed with DISTRUCT 1.1 (Rosenberg 2004). The strongest level of the genetic subdivision was determined using ΔK (Evanno et al. 2005), as implemented in the online post-processing software Structure Harvester (http://taylor0.biology.ucla.edu/structureHarvester/) (Earl and vonHoldt 2012). Principal components analyses (PCA) were performed to investigate the genetic structure of *P. brigantina* using the scatterplot3d R package (Ligges and Mächler 2003) or among the five *Prunophora* species, using the DARwin software package v 6.0.017 (Perrier and Jacquemoud-Collet 2006). Further genetic differentiation and relationships were also estimated using a weighted neighbour-joining tree as implemented in the DARwin software package v 6.0.017 (Perrier and Jacquemoud-Collet 2006).

We performed a three-step population subdivision analysis, the first one with only *P. brigantina* samples, and the second one adding previously obtained datasets of *Armeniaca* species, including *P. armeniaca*, *P. sibirica*, *P. mandshurica* and *P. mume* wild and cultivated samples (Liu et al. 2019) (dataset 2 in Table 1). In the third step of the analysis, we added samples of the Prunus (*P. cerasifera* and *P. salicina*) and Prunocerasus (*P. mexicana*, *P. munsoniana*, *P. maritima*, *P. americana* and *P. subcordata)* sections (dataset 3 in Table 1). The same procedure to investigate population subdivision and structure analysis was used with the Armeniaca, Prunus and Prunocerasus diploid samples. In parallel, we also performed a population subdivision analysis along the native distribution of *P. cerasifera* (Supplemental information '*Prunus cerasifera* diversity and population structure analysis' and dataset 4 in Table 1). In *P. cerasifera*, a Neighbour-Joining tree based on Nei's standard genetic distance was built with a bootstrap of 30,000 in PopTreeW (Takezaki et al. 2014).

In order to test whether there was a pattern of isolation by distance, we performed a Mantel test between a matrix of Edwards' genetic distances and a matrix of Euclidean

226 geographic distances in *P. brigantina* using the R adegenet package (Jombart and Ahmed
227 2011).

228

Genetic diversity, differentiation and core collection constitution

230 We used GENALEX 6.501 (Peakall and Smouse 2012) to estimate the number of alleles
231 (N_a), the effective number of alleles (N_e), i.e., the number of equally frequent alleles that
232 would achieve the same expected heterozygosity as in the sample, the observed
233 heterozygosity (H_O), the unbiased expected heterozygosity (H_E) and the Shannon index
234 (*I*) (Shannon 1948). Genetic differentiation among genetic clusters (Jost′D) was estimated
235 in GENODIVE (Meirmans and Van Tienderen 2004). The allelic richness (A_r) and the
236 private allelic richness (A_p) were calculated after adjustment for sample size differences
237 among groups through the rarefaction procedure implemented in ADZE Allelic Diversity
238 Analyzer v1.0 (Szpiech et al. 2008), setting the sample size to five.

239　　　The maximization (M) strategy (Schoen and Brown 1993) implemented in the
240 COREFINDER software was used to generate a core *P. brigantina* tree collection
241 maximizing the number of alleles based on our dataset. The maximization strategy
242 consisted in detecting the smallest sample that captured 100% of the genetic diversity
243 present within the entire germplasm collection. We further used the Mann-Whitney U test
244 to check the genetic diversity difference between the core collection and the entire *P.*
245 *brigantina* sample.

246

Results

Genetic diversity and population structuration in *P. brigantina*

249 Thirty-four microsatellite markers used in a previous study (Liu et al. 2019) were tested
250 for our *P. brigantina* population study. Four markers (AMPA109, ssr02iso4G, BPPCT008
251 and BPPCT038) failed to amplify or generated over 50% of missing data and were
252 consequently eliminated. Six other markers (BPPCT030, CPPCT022, CPSCT004,
253 UDP98-412, UDA-002 and PacB26) gave poor amplification in *P. brigantina*, yielding
254 more than 10% missing data. This may be because of poor marker transferability, as most
255 of the above microsatellite markers were developed from genomic data on other *Prunus*
256 species, such as peach, almond, apricot and Japanese plum. The remaining 24
257 microsatellite markers performed well in *P. brigantina* and were used in this study (Table
258 S2). In our *P. brigantina* sample, the number of alleles (N_A) was 121 (mean of 5.04 per

259 marker) and the number of effective alleles (N_E) 59.57 (mean of 2.48 per marker) (Table
260 S2).

261   The biologically most relevant genetic clustering of *P. brigantina* was found to be
262 *K*=3: the DeltaK statistics indicated that it was the strongest population subdivision level
263 (Figure S1A) and further increasing *K* yielded many admixed individuals (Figure S2). The
264 three inferred genetic clusters (blue, yellow and orange colours in Figure 2b and S2)
265 corresponded to three French national parks "Queyras", "Ecrins" and "Mercantour",
266 respectively. Weak but significant genetic differentiation (mean Jost's *D*=0.117) was found
267 among these three *P. brigantina* populations (Table 2). Both the Josts' *D* and the PCA
268 indicated that the *P. brigantina* "Queyras" cluster was the most differentiated from the two
269 other ones, the "Ecrins" and "Mercantour" clusters being found genetically closer one to
270 each other (Table 2, Figure 3).

271   An additional subdivision of the Ecrins cluster was found at *K*=8, revealing
272 differentiation between the dark blue and yellow clusters (Figure S2). The Mantel test on
273 the three *P. brigantina* clusters indicated no significant relationship between genetic
274 differentiation and geographic distance (*P*=0.308, by Monte Carlo permutation tests,
275 Figure S3), indicating a lack of isolation by distance.

276

277 **Genetic relationships between *P. brigantina* and other Prunophora species**

278 To obtain a better understanding of the genetic relationships between *P. brigantina* and
279 other Armeniaca species, we combined the current *P. brigantina* data with a former
280 Armeniaca dataset built with the same 24 microsatellite markers (Liu et al. 2019) (Tables
281 1 and S1c). We performed a Bayesian clustering analysis on the full Armeniaca dataset,
282 including wild and cultivated *P. armeniaca, P. sibirica, P. mandshurica* and *P. mume*
283 (Table 1). We obtained a similar structure as the one described in (Liu et al. 2019) for the
284 previous dataset, and *P. brigantina* differentiated in a distinct cluster, from *K*=3 and above
285 (yellow colour in Figure S4). No gene flow with other species of the section Armeniaca
286 was detected (i.e. no individuals who would have admixed ancestry between the yellow
287 cluster and other clusters), in particular in between wild *P. brigantina* and cultivated
288 apricots which, yet, partly share habitats over Western Europe (Figure S4).

289   We further questioned the genetic relationship of *P. brigantina* with other members
290 of the Prunophora subgenus, i.e. species of the Prunus and Prunocerasus sections.
291 Because *P. cerasifera* (cherry plum), a species of the Prunus section, is partly sharing
292 habitats with *P. brigantina*, we significantly extended the sampling of *P. cerasifera* species
293 compared to (Horvath et al. 2008), including accessions from the cherry plum native area,

9

294   i.e. Caucasia and Central Asia, to obtain a better representation at the species level. We

295   then explored the genetic differentiation of this species over its Eurasian distribution. We

296   found genetically differentiated clusters of cherry plums, with contrasted geographical

297   distributions from Central Asia to Europe (detailed results are presented in the

298   supplemental information '*Prunus cerasifera* diversity and population structure analysis).

299   Caucasia appears to be a diversification center of wild cherry plums, with two distinct

300   genetic clusters that may result from geographical isolation. This dataset was later merged

301   with representatives of the Prunus, Armeniaca and Prunocerasus sections, to infer the

302   origin of *P. brigantina* and its genetic relationships with species of the Prunophora

303   subgenus (Tables 1, S1b and S1c). In the following step, we focused on species that

304   shared, partly, habitats with *P. brigantina*, i.e. *P. cerasifera* and cultivated *P. armeniaca*,

305   together with other Armeniaca (*P. mume*), Prunus (*P. salicina*) and Prunocerasus species.

306   For this, we used genotyping data based on 23 microsatellite markers (see the

307   supplemental information '*Prunus cerasifera* diversity and population structure analysis').

308   The delta *K* peaked at *K*=3, indicating that this was the strongest level of population

309   subdivision (Figure S1B). However, further relevant clustering was observed at higher *K*

310   values (Figure S5). From *K*=7 and above, all taxonomic species separated in specific

311   clusters: green for *P. brigantina*, pink for *P. armeniaca*, blue for *P. cerasifera,* grey for *P.*

312   *mume*, orange for *P. salicina* and black for Prunocerasus (Figure S5). Again, we could not

313   find any admixture footprints between *P. brigantina* and other Prunus species, while there

314   may be some footprints of introgression from *P. cerasifera* into *P. salicina* (see admixed

315   individuals indicated by blue stars in Figure S5), although the blue and orange

316   heterogeneous bars may alternatively result from low assignment power due to the low

317   number of *P. salicina* individuals.

318   We further explored the genetic differentiation and relationships among all

319   Prunophora samples using an unrooted weighted neighbour-joining tree (Figure 4). In the

320   tree, the delimitation of *P. brigantina* as a distinct species from other apricot and plum

321   taxonomic species was well supported (100% bootstrap support). *Prunus brigantina* trees

322   appeared genetically closer to the Armeniaca species (*P. armeniaca* and *P. mume*) than

323   to other Prunus and Prunocerasus species, which is consistent with Rehder's taxonomy.

324   The principal component analysis (PCA) supported the differentiation of *P. brigantina* from

325   other species of the Armeniaca section, and from the Prunus and Prunocerasus sections

326   (Figure 5). Both the NJ tree and the PCA indicated that plum species (*P. cerasifera* and

327   *P. salicina*) were partly overlapping, in particular the cultivated Japanese plums and cherry

328   plums; the wild *P. salicina* trees in contrast appeared well separated from *P. cerasifera*

329   (Figures 4 and 5). The overlapping may be the result of low power to distinguish the groups

10

330 based on few individuals or of hybridization between cultivated trees. One particular case

331 of hybridization is P0489, cv. Rutland plumcot. Breeders' information indicates that it is a

332 hybrid between plum and apricot. In our structure barplots, NJ tree and PCA (Figures S5,

333 4 and 5), P0489 in fact appeared admixed between the two plum species, *P. salicina* and

334 *P. cerasifera,* and not with apricot.

335

Construction of a *P. brigantina* core collection

337   We used the COREFINDER program to identify the smallest core collection that

338 would be sufficient to capture the whole diversity detected based on our 24 microsatellite

339 markers. Based on the maximizing strategy implemented in COREFINDER, we propose

340 the use of a core set of 36 individuals (~49% of the whole *P. brigantina* sample) that

341 captures 100% of the detected diversity (Figure 6, Table S3). Pairwise comparisons using

342 Mann-Whitney U tests showed no significant differences in diversity indexes *(I, H_O,* and

343 *H_E)* between the *P. brigantina* entire Alpine sample (*N*=71) and the core collection (*N*=36)

344 (Tables S2, S4 and S5). This indicates that our core collection can be used as an *ex-situ*

345 germplasm repository.

346

Discussion

348   The current study showed that *P. brigantina* is still found in a few Alpine valleys,

349 along the border between France and northwest Italy, where it grows above 1,400 m

350 altitude as single isolated trees (except for the plateau of Nevache, where they are present

351 as a denser population), in arid places such as shrub thickets. In France, it is confined to

352 the three southeastern departments of Alpes-Maritimes, Alpes-de-Haute-Provence and

353 Hautes-Alpes. The sustainability of *P. brigantina* habitat is threatened by forest

354 fragmentation. This raises the question of the long-term conservation of this species and

355 no germplasm accessions of *P. brigantina* are reported by EURISCO to be held in

356 European ex-situ collections. Because large field collections of perennial crops are

357 expensive to maintain, the identification of a restricted number of representatives of *P.

358 brigantina* population for *ex situ* conservation would be very useful in the perspective of

359 Alpine ecosystem restauration and future breeding programs. Core collections are

360 representative subsets of germplasm collections that are developed to improve the

361 efficiency of germplasm evaluation while increasing the probability of finding genes of

362 interest (Simon and Hannan 1995). Therefore, our current core collection will serve in the

363 future for *P. brigantina* conservation as well as for stone fruit breeding programs benefiting

364 from *P. brigantina* resilience characteristics, especially in a context of Mediterranean

365 climate changes. However, the most efficient strategy for biodiversity conservation
366 remains the preservation of the natural habitat of endangered species.

367   Thanks to an extensive dataset of Prunophora species, we also questioned here
368 the genetic relationships of *P. brigantina* with other species of the Prunus and Armeniaca
369 sections. Species of the Prunocerasus section were not integrated in the analysis except
370 as outgroups because they are naturally distributed on different continents and do not
371 overlap in their respective natural habitats with *P. brigantina*. Through Bayesian analyses,
372 *P. brigantina* appears as a *bona fide* species, clearly distinct from other apricot species
373 and from plum species, with no footprint of admixture. Our results are in accordance with
374 previous studies that indicate a clear differentiation of *P. brigantina* from other Armeniaca
375 apricot species (i.e. *P. armeniaca* and *P. mume*) but do not support a close relationship
376 with species of the Prunus section (Chin et al. 2014; Reales et al. 2010; Shi et al. 2013;
377 Zhebentyayeva et al. 2019). This might be due to the fact that our sampling covers a larger
378 diversity panel than in the former studies, both in *Armeniaca* and *Prunus* sections, *P.*
379 *brigantina* included. Indeed, sampling only one or two individuals per species is expected
380 to lower the accuracy of phylogenetic analyses (Wiens and Servedio 1997). In our
381 analyses, *P. brigantina* was closer to species of the Armeniaca section than to the Prunus
382 section. While *P. brigantina* should still be considered as an Armeniaca species, it has
383 diverged from *P. armeniaca* long before *P. mume*, thus representing the most genetically
384 distant apricot-related species within the Armeniaca section (Hagen et al. 2002; Liu et al.
385 2019).

386   Contradictory results had been obtained from a phylogeny of Eurasian plum
387 species based on chloroplast DNA sequences (Reales et al. 2010), where *P. brigantina*
388 grouped together with European Prunus species, such as the polyploid *P. spinosa*, *P.*
389 *insititia* and *P. domestica*, and the diploid *P. ramburii* Boiss.species; it was clearly
390 separated from *P. armeniaca* (apricot). The proximity in chloroplast genotypes between
391 *P. brigantina* and the polyploid Prunus species might indicate the Alpine plum as a
392 parental contributor in interspecific hybridization of polyploid Prunus species
393 (Zhebentyayeva et al. 2019). Organelles are however known to introgress much more
394 often than nuclear DNA and chloroplast genealogies are often discordant from nuclear
395 phylogenies (Coyne and Orr 2004).The other plum species that grouped with *P. brigantina*
396 in chloroplast genealogy, *P. ramburii,* is a relict, wild species endemic in the southern
397 Spanish mountains (Sierra Nevada and Sierra Baz) While its distribution is in Europe, it
398 does not overlap with that of *P. brigantina.* Hence, its morphological features are distinct
399 from Alpine plum, forming bushes with tiny, blue/violet drupes and narrow leaves
400 (http://www.anthos.es/index.php?lang=en). Therefore, the incongruence between our

results with those obtained earlier based on the chloroplast genome echoes the conclusions of others that despite the many advantages and widespread use of chloroplast DNA in phylogenetic studies, caution has to be taken in the use of organellar variation for inferring phylogeny (Doyle 1992; Lee-Yaw et al. 2019; Soltis and Kuzoff 1995).

Nevertheless, by extending the sampling set of both *P. brigantina* and plum species, our study provides compelling evidence that *P. brigantina* grouped in the Armeniaca section. It illustrates the importance of the sample size and sampling design that encompasses here a larger genetic diversity at the species level than in previous studies (Hagen et al. 2002; Horvath et al. 2008; Reales et al. 2010; Zhebentyayeva et al. 2019). It also questioned the relevance of the classification into sections of the Prunophora subgenus, at least for the Eurasian sections, i.e. Armeniaca and Prunus. Species of the two sections are sharing habitats and they are interfertile, in particular between diploid species, thus resulting in a number of hybrids and probably new species (Cici and Van Acker 2010; Layne and Sherman 1986). Although the genetic differentiation of the Prunus and Armeniaca sections from the Prunocerasus section is clear (Krüssmann 1978), the relationships among taxa of the two Eurasian sections are not well resolved as illustrated by the role of cross taxa hybridization in Japanese apricot (*P. mume*) adaptive evolution (Numaguchi et al. 2020). The previous controversial classification of *P. brigantina* either in the Armeniaca section or in the Prunus section reflects the difficulty of assigning a clear barrier between species of those two sections; an analysis of the entire subgenus using a shared set of same nuclear markers could provide greater resolution and would place the findings presented here into a Prunophora-wide perspective.

Conclusion

In this study, we found a low level of genetic diversity in natural *P. brigantina* populations and identified three genetically differentiated populations, in the Ecrins, Queyras and Mercantour national parks, respectively. We further successfully established in Bordeaux a core collection of 36 individuals representing the *P. brigantina* diversity that will be publicly available through the French Genetic Resources Center. In addition, a population NJ tree did not support a close relationship between *P. brigantina* and the other Prunus species, *P. brigantina* being closer to Armeniaca species whilst remaining clearly distinct. While most of the fruit species originate from Asia or America, many crop wild relatives still exist both in their center of origin and along their dispersal routes. For example, in pit and stone fruits, several *Prunus, Malus* and *Pyrus* wild species are endemic in Europe

436 and are often threatened by the rapid changes of land use (Cornille et al. 2013a; Welk et

437 al. 2016). To inform *in situ* and *ex situ* conservation measures and add value to fruit tree

438 genetic resources, we recommend in-depth characterization of those wild relatives,

439 similarly to the current study in the Alpine plum.

440

451

452 **Statements**

453 Appropriate permissions from responsible authorities for collecting and using *Prunus*

454 samples from Central Asia and Caucasia were obtained by the local collaborators. The

455 official authorization for the survey and sampling of *P. brigantina* genetic resources is

456 registered and accessible through the following link:

457 https://absch.cbd.int/database/ABSCH-IRCC-FR-246978. The rest of the samples were

458 kindly provided, with due authorizations, by the curators of the French INRAE Genetic

459 Resources Centre (GRC) and the US ARS-USDA repository, further details are available

460 on their respective databases.

461

462 **Data availability**

463 The datasets generated by the current study, *i.e*. the SSR genotyping, are available at

464 the INRAE data portal (https://data.inrae.fr/) where they can be freely retrieved.

465

466 ***Captions for the supplementary Figures presented in a separate PDF file***

467 **Figure S1. DeltaK plot as a function of K for the *Prunus brigantina* (A) and**

468 **Prunophora (B) dataset.**

469 **Figure S2. Bayesian clustering on *Prunus brigantina* samples in the French Alps.**

470 *Prunus brigantina* dataset included 71 individuals sampled from the French Alps and two

471 samples from the French GRC repository. Each individual is represented by a vertical bar,

472 partitioned into *K* segments representing the inferred proportions of ancestry of its

473 genome.

474 **Figure S3. Isolation by distance (IBD) test in *Prunus brigantina*.**

475 a. Distribution of correlation values between genetic and geographic distances under the

476 assumption of lack of isolation by distance, drawn from permutations; the observed value

477 of the correlation between the distance matrices, represented by the black diamond, falls

478 within the expected distribution which indicates the lack of isolation by distance pattern.

479 b. Pairwise Edwards' distances plotted against Euclidean geographic distances, with local

480 density of points plotted using a two-dimensional kernel density estimate, displayed in

481 colour from white to red. The solid line represents the fitted linear regression between

482 Edwards' genetic and Euclidean geographic distances.

483 **Figure S4. Bayesian analysis on Armeniaca and wild *Prunus brigantina***

484 **accessions.**

485 Genetic subdivision among Armeniaca species, *P. brigantina* included, was inferred with

486 STRUCTURE with 24 microsatellite markers. The 648 samples belong to the six

487 Armeniaca species as follows: *P. brigantina* (*N*=73), *P. armeniaca* (European and Chinese

488 cultivated *N*=270 and wild, *N*=204), *P. sibirica* (*N*=84), *P. mume* (*N*=9), *P. mandshurica*

489 (*N*=8). Each individual is represented by a vertical bar, partitioned into *K* segments

490 representing the inferred proportions of ancestry of its genome. Species and origin of the

491 accessions are indicated on the top of the figure.

492 **Figure S5. Bayesian analysis on the *Prunus brigantina* dataset together with an**

493 **extended Prunophora dataset.**

494 Genetic subdivision among Armeniaca, Prunus and Prunocerasus species was inferred

495 with STRUCTURE with 23 microsatellite markers (supplemental information for the list of

496 markers). The 226 samples belong to three different Prunophora species including *P.*

497 *brigantina* (*N*=73), *P. cerasifera* (*N*=66), *P. armeniaca* (*N*=87), *P. salicina* (*N*=10), *P.*

498 *mume* (*N*=9), *P. mexicana* (*N*=1), *P. munsoniana* (*N*=1), *P. maritima* (*N*=1), *P. americana*

499 (*N*=1) and *P. subcordata* (*N*=1). The blue stars (*), at the bottom of the bar plots,

500 correspond to Japanese plums (*P. salicina*) admixed with *P. cerasifera*.

501

502 *Legends for the supplementary tables presented in a separate PDF file*

503 **Table S1a. Sampling locations, geographic regions and assigned genetic cluster of**

504 ***Prunus brigantina* in the French Alps.**

505 FR for an origin from the French Alps. Sampling site is indicated in GPS coordinates, N

506 for North, E for East.

507 **Table S1b. Sampling locations, geographic regions and/or germplasm repositories**

508 **of *Prunus cerasifera* samples.**

509 [1] Species affiliation as indicated by the curator of the germplasm collection where the

510 sample is maintained or as identified *in situ*. [2] Sampling location in decimal degrees. [3]

511 Origin as indicated in the database of the germplasm repository. n/a, not applicable

512 because admixed and thus not used in the correlation tests

513 **Table S1c: List of individuals included in the different datasets.**

514 [1] FR refers to France, AZ to Azerbaijan, CH to China, KR to Kyrgyzstan, KZ to Kazakhstan,

515 OUZ to Uzbekistan, TCH to Czech republic (Lednice repository), TURC to Turkey

516 (Malatya repository), US to USA (ARS-USDA Prunus germplasm repository). For more

517 details, see Liu et al (2019). Accession numbers starting with A indicate apricot cultivars

518 and with P, plum cultivars, as displayed in the French GRC database. The sign (-) means

519 that the sample is maintained in germplasm repository and was not collected *in situ*. The

520 cross in the last four columns (dataset 1 to 4) means that this sample was used in the

521 corresponding dataset.

522 **Table S2. Analysis of genetic variability from microsatellite markers for *Prunus***

523 ***brigantina* population.**

524 *Na*: number of different alleles, and *Ne*: number of effective alleles. *I*: Shannon diversity

525 index. *He* and *Ho*: expected and observed heterozygosities.

526 **Table S3. The description of individuals retained for the core collection of *Prunus***

527 ***brigantina***

528 **Table S4. Genetic variability of microsatellite markers for the *Prunus brigantina***

529 **core collection.**

530 *Na*: number of different alleles, and *Ne*: number of effective alleles. *I*: Shannon diversity

531 index. He and *Ho*: expected and observed heterozygosities.

16

532 **Table S5. Mann-Whitney U tests (two-tailed) between the whole *Prunus brigantina***

533 **dataset and its core collection.**

534

535 ***Supplemental information presented in a separate PDF file***

536 **Supplemental information '*Prunus cerasifera* diversity and population structure**

537 **analysis'**

538

539 **References**

540 Allendorf FW, Luikart GH, Aitken SN (2012) Conservation and the Genetics of Populations.,
541   2nd Edition edn. Wiley-Blackwell Ltd,
542 Branca F, Donnini D (2011) *Prunus brigantina.* vol e.T172164A121228349. .
543   doi:http://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T172164A6840507.en.
544 Chin S-W, Shaw J, Haberle R, Wen J, Potter D (2014) Diversification of almonds, peaches,
545   plums and cherries – Molecular systematics and biogeographic history of *Prunus*
546   (Rosaceae). Molecular Phylogenetics and Evolution 76:34-48
547   doi:https://doi.org/10.1016/j.ympev.2014.02.024
548 Cici S, Van Acker R (2010) Gene flow in *Prunus* species in the context of novel trait risk
549   assessment. Environmental biosafety research 9:75-85 doi:10.1051/ebr/2010011
550 Cornille A et al. (2013a) Postglacial recolonization history of the European crabapple (*Malus*
551   *sylvestris* Mill.), a wild contributor to the domesticated apple. Molecular Ecology
552   22:2249-2263 doi:10.1111/mec.12231
553 Cornille A, Gladieux P, Giraud T (2013b) Crop-to-wild gene flow and spatial genetic structure in
554   the closest wild relatives of the cultivated apple. Evolutionary Applications 6:737-748
555   doi:10.1111/eva.12059
556 Couplan F (2009) Le régal végétal: plantes sauvages comestibles. Collection L'encyclopédie des
557   plantes sauvages comestibles et toxiques de l'Europe.
558 Coyne JA, Orr HA (2004) Speciation. OUP USA,
559 Decroocq S et al. (2016) New insights into the history of domesticated and wild apricots and its
560   contribution to *Plum pox virus* resistance. Molecular Ecology 25:4712-4729
561   doi:10.1111/mec.13772
562 Doyle JJ (1992) Gene Trees and Species Trees: Molecular Systematics as One-Character
563   Taxonomy. Systematic Botany 17:144-163 doi:10.2307/2419070
564 Dupouy J (1959) Le prunier de Briançon ou Marmottier (*Prunus brigantiaca* Vill.) et les huiles
565   de marmotte. Bull SAJA 141:69-71
566 Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for
567   visualizing STRUCTURE output and implementing the Evanno method. Conservation
568   Genetics Resources 4:359-361 doi:10.1007/s12686-011-9548-7
569 Escribano P, Viruel M, Hormaza J (2008) Comparison of different methods to construct a core
570   germplasm collection in woody perennial species with simple sequence repeat markers. A
571   case study in cherimoya (*Annona cherimola*, Annonaceae), an underutilised subtropical
572   fruit tree species. Annals of Applied Biology 153:25-32 doi:10.1111/j.1744-
573   7348.2008.00232.x
574 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the
575   software structure: a simulation study. Molecular Ecology 14:2611-2620
576   doi:10.1111/j.1365-294X.2005.02553.x

577 Fahrig L (2003) Effects of Habitat Fragmentation on Biodiversity. Annual Review of Ecology,
578   Evolution, and Systematics 34:487-515 doi:10.1146/annurev.ecolsys.34.011802.132419
579 Frankel OH, Brown AHD Current Plant genetic resources today - a critical appraisal. In:
580   Proceedings of the XV International Congress of Genetics. Crop Genetic Resources:
581   Conservation and Evaluation, New Delhi, India, 1983. Oxford & IBH publishing CO., pp
582   241-256
583 Glaszmann JC, Kilian B, Upadhyaya HD, Varshney RK (2010) Accessing genetic diversity for
584   crop improvement. Current Opinion in Plant Biology 13:167-173
585   doi:10.1016/j.pbi.2010.01.004
586 Govindaraj M, Vetriventhan M, M S (2014) Importance of Genetic Diversity Assessment in Crop
587   Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. Genetics
588   Research International 2015 doi:10.1155/2015/431487
589 Hagen L, Khadari B, Lambert P, Audergon J-M (2002) Genetic diversity in apricot revealed by
590   AFLP markers: species and cultivar comparisons. Theoretical and Applied Genetics
591   105:298-305 doi:10.1007/s00122-002-0910-8
592 Heled J, Drummond AJ (2010) Bayesian Inference of Species Trees from Multilocus Data.
593   Molecular Biology and Evolution 27:570-580 doi:10.1093/molbev/msp274
594 Horvath A, Christmann H, Laigret F (2008) Genetic diversity and relationships among *Prunus*
595   *cerasifera* (cherry plum) clones. Botany 86:1311-1318 doi:10.1139/b08-097
596 Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for
597   dealing with label switching and multimodality in analysis of population structure.
598   Bioinformatics 23:1801-1806 doi:10.1093/bioinformatics/btm233
599 Jombart T, Ahmed I (2011) Adegenet 1.3-1: new tools for the analysis of genome-wide SNP
600   data. Bioinformatics 27:3070-3071 doi:10.1093/bioinformatics/btr521
601 Krüssmann G (1978) Manual of cultivated broad-leaved trees and shrubs. vol 3 Pru-Z. Timber
602   Press, Portland
603 Layne RE, Sherman CWB (1986) Interspecific hybridization of Prunus. HortScience 21:48-51
604 Lee-Yaw JA, Grassa CJ, Joly S, Andrew RL, Rieseberg LH (2019) An evaluation of alternative
605   explanations for widespread cytonuclear discordance in annual sunflowers (Helianthus).
606   New Phytologist 221:515-526 doi:10.1111/nph.15386
607 Li DZ, Pritchard HW (2009) The science and economics of ex situ plant conservation. Trends in
608   plant science 14:614-621
609 Ligges U, Mächler M (2003) Scatterplot3d - an R package for Visualizing Multivariate Data.
610   Journal of Statistical Software 8:1-20 doi:10.17877/de290r-15194
611 Liu S et al. (2019) The complex evolutionary history of apricots: Species divergence, gene flow
612   and multiple domestication events. Molecular Ecology 28:5299-5314
613   doi:10.1111/mec.15296
614 Meirmans PG, Van Tienderen PH (2004) Genotype and genodive: two programs for the analysis
615   of genetic diversity of asexual organisms. Molecular Ecology Notes 4:792-794
616   doi:10.1111/j.1471-8286.2004.00770.x
617 Noble V, Van Es J, Michaud H, Garraud L (2015) Liste Rouge de la flore vasculaire de
618   Provence-Alpes-Côte d'Azur. Provence-Alpes-Côte d'Azur, France
619 Numaguchi K, Akagi T, Kitamura Y, Ishikawa R, Ishii T (2020) Interspecific introgression and
620   natural selection in the evolution of Japanese apricot (*Prunus mume*).
621   bioRxiv:2020.2006.2023.141200 doi:10.1101/2020.06.23.141200
622 Peakall R, Smouse P (2012) GenAlEx V5: Genetic Analysis in Excel. Populations Genetic
623   Software for Teaching and Research. Bioinformatics (Oxford, England) 28:2537-2539
624   doi:10.1093/bioinformatics/bts460
625 Perrier X, Jacquemoud-Collet JP (2006) DARwin software. http://darwin.cirad.fr/.
626 Pignatti S (1982) Flora d'Italia. vol 2. Edagricole, Bologna, Italy
627 Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus
628   Genotype Data. Genetics 155:945-959

629 Reales A, Sargent DJ, Tobutt KR, Rivera D (2010) Phylogenetics of Eurasian plums, Prunus L.
630   section Prunus (Rosaceae), according to coding and non-coding chloroplast DNA
631   sequences. Tree Genetics & Genomes 6:37-45 doi:10.1007/s11295-009-0226-9
632 Rehder A (1940) Manual of cultivated trees and shrubs hardy in North America. vol 1.
633   Biosystematics, floristic & phylogeny series. Collier Macmillan Ltd, New York
634 Rosenberg NA (2004) Distruct: a program for the graphical display of population structure.
635   Molecular Ecology Notes 4:137-138 doi:10.1046/j.1471-8286.2003.00566.x
636 Schoen DJ, Brown AHD (1993) Conservation of allelic richness in wild crop relatives is aided by
637   assessment of genetic markers. Proceedings of the National Academy of Sciences
638   90:10623-10627 doi:10.1073/pnas.90.22.10623
639 Shannon CE (1948) A Mathematical Theory of Communication. Bell System Technical Journal
640   27:379-423 doi:10.1002/j.1538-7305.1948.tb01338.x
641 Shi S, Li J, Sun J, Yu J, Zhou S (2013) Phylogeny and Classification of Prunus sensu lato
642   (Rosaceae). Journal of Integrative Plant Biology 55:1069-1079 doi:10.1111/jipb.12095
643 Simon CJ, Hannan RM (1995) Development and Use of Core Subsets of Cool-season Food
644   Legume Germplasm Collections. HortScience 30:907C doi:10.21273/hortsci.30.4.907c
645 Soltis DE, Kuzoff RK (1995) Discordance between nuclear and chloroplast phylogenies in the
646   Heuchera group (Saxifragaceae). Evolution 49:727-742 doi:10.1111/j.1558-
647   5646.1995.tb02309.x
648 Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting
649   alleles private to combinations of populations. Bioinformatics 24:2498-2504
650   doi:10.1093/bioinformatics/btn478
651 Takeda T, Shimada T, Nomura K, Ozaki T, Haji T, Yamaguchi M, Yoshida M (1998)
652   Classification of Apricot Varieties by RAPD Analysis. Engei Gakkai Zasshi 67:21-27
653   doi:10.2503/jjshs.67.21
654 Takezaki N, Nei M, Tamura K (2014) POPTREEW: web version of POPTREE for constructing
655   population trees from allele frequency data and computing some other quantities. Mol
656   Biol Evol 31:1622-1624 doi:10.1093/molbev/msu093
657 Tison J-M, De Foucault B (2014) Flora Gallica - Flore de France. Biotope, Mèze
658 Villars D (1786) Histoire des plantes de Dauphiné: Contenant une Préface Historique, un
659   Dictionnaire des Termes de Botanique, les Classes, les Familles, les Genres, & les
660   Herborisations des Environs de Grenoble, de la Grande Chartreuse, de Briançon, de Gap
661   & de Montelimar. vol 1. Prevost, Paris
662 Welk E, de Rigo D, Caudullo G (2016) *Prunus avium* in Europe: distribution, habitat, usage and
663   threats. In: San-Miguel-Ayanz J, de Rigo D, Caudullo G, Houston Durrant T, A. M (eds)
664   European Atlas of Forest Tree Species. Publication Office of the European Union,
665   Luxembourg,
666 Wiens JJ, Servedio MR (1997) Accuracy of Phylogenetic Analysis Including and Excluding
667   Polymorphic Characters. Systematic Biology 46:332-345 doi:10.1093/sysbio/46.2.332
668 Zhebentyayeva T et al. (2019) Genetic characterization of worldwide *Prunus domestica* (plum)
669   germplasm using sequence-based genotyping. Horticulture Research 6
670   doi:10.1038/s41438-018-0090-6
671
672

Figures

Figure 1. Taxonomy and geographic distribution of the different species in the Armeniaca section. Species classification is based on reports by Rehder (1940). Data on species distribution were retrieved from the global biodiversity information facility (GBIF) (https://doi.org/10.15468/39omei). Dots represent georeferenced species records from 1910 to 2017.
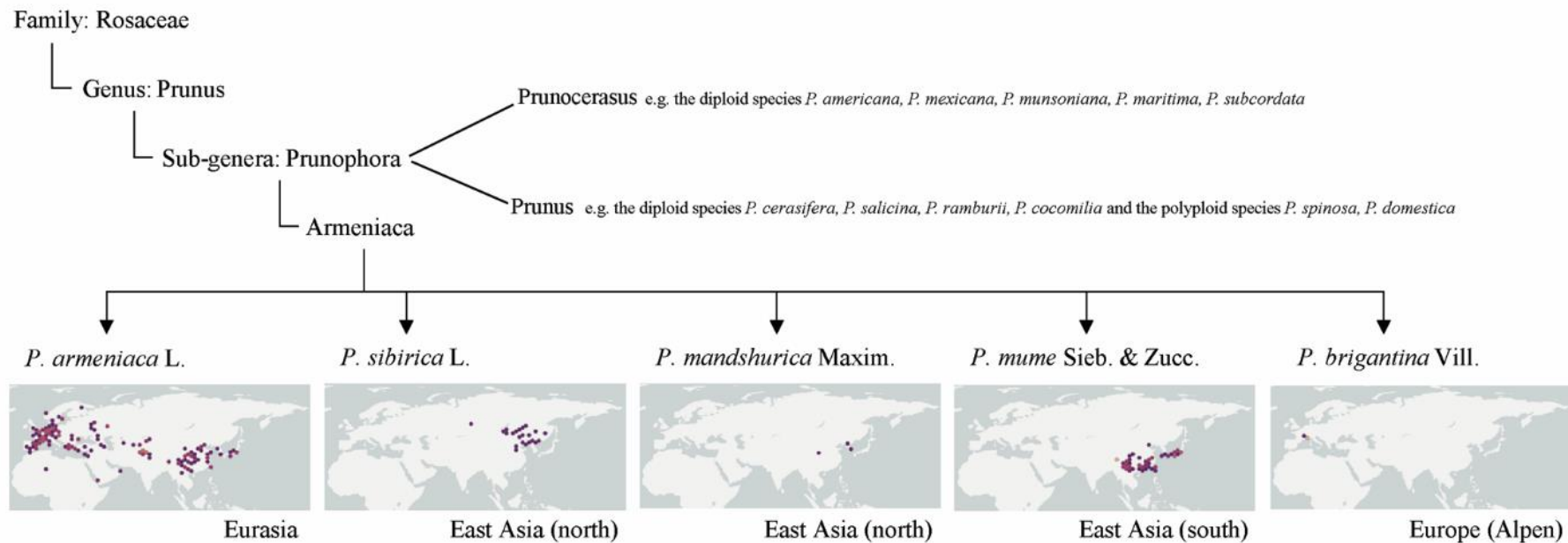
681 **Figure 2. *Prunus brigantina* morphological features, genetic clustering and spatial**

682 **distribution in the French Alps.** a. A *P. brigantina* small tree in its natural habitat

683 (Arvieux) (left), summery leaves (middle) and ripening fruits (right). b. The three genetic

684 clusters of *P. brigantina* inferred from the STRUCTURE analysis (Figure S2 at *K*=3) and

685 their spatial distribution in the French Alps. "Ecrins", "Queyras" and "Mercantour" refer to

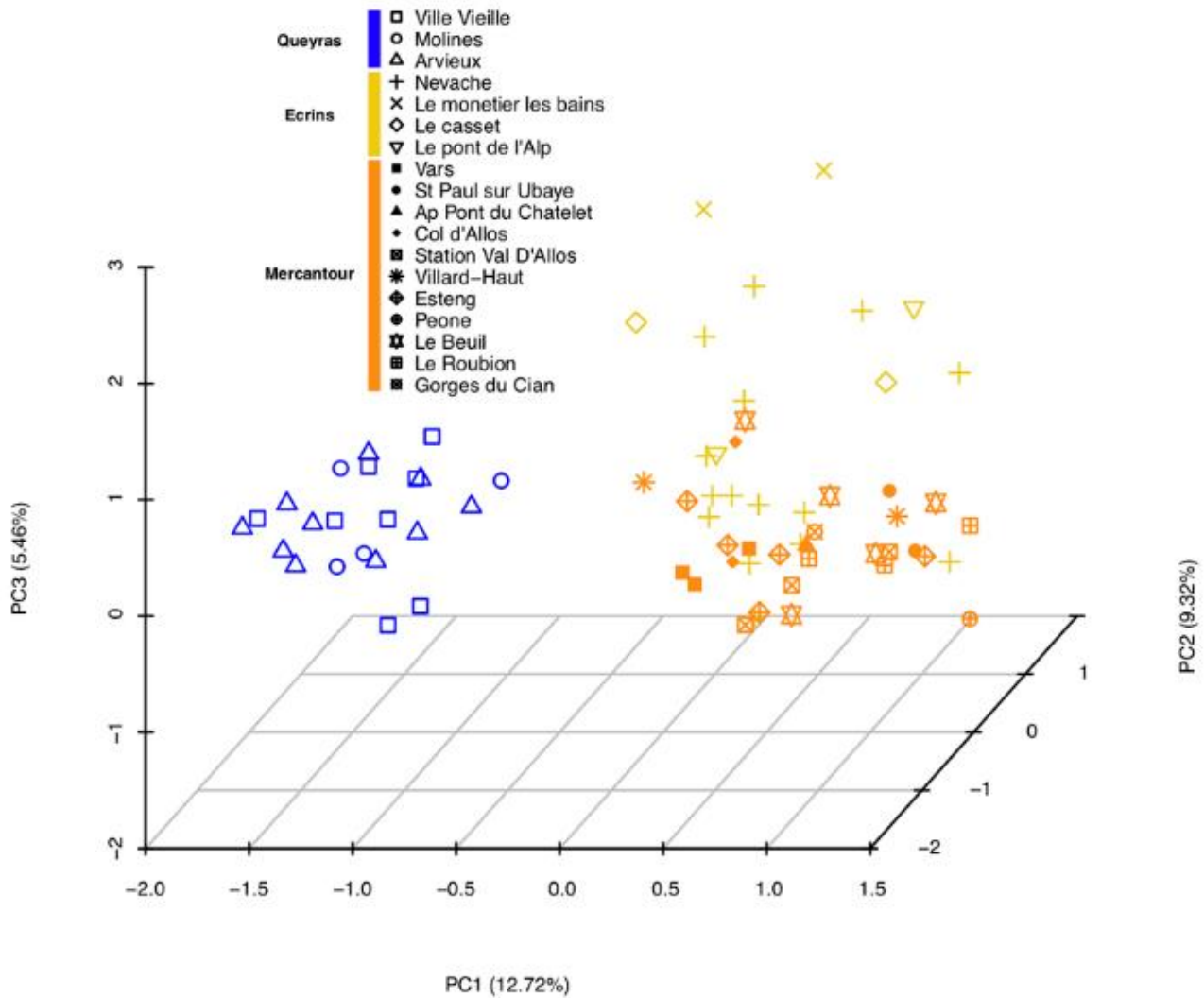686 the three national parks in the southeast of France.

687
688

21
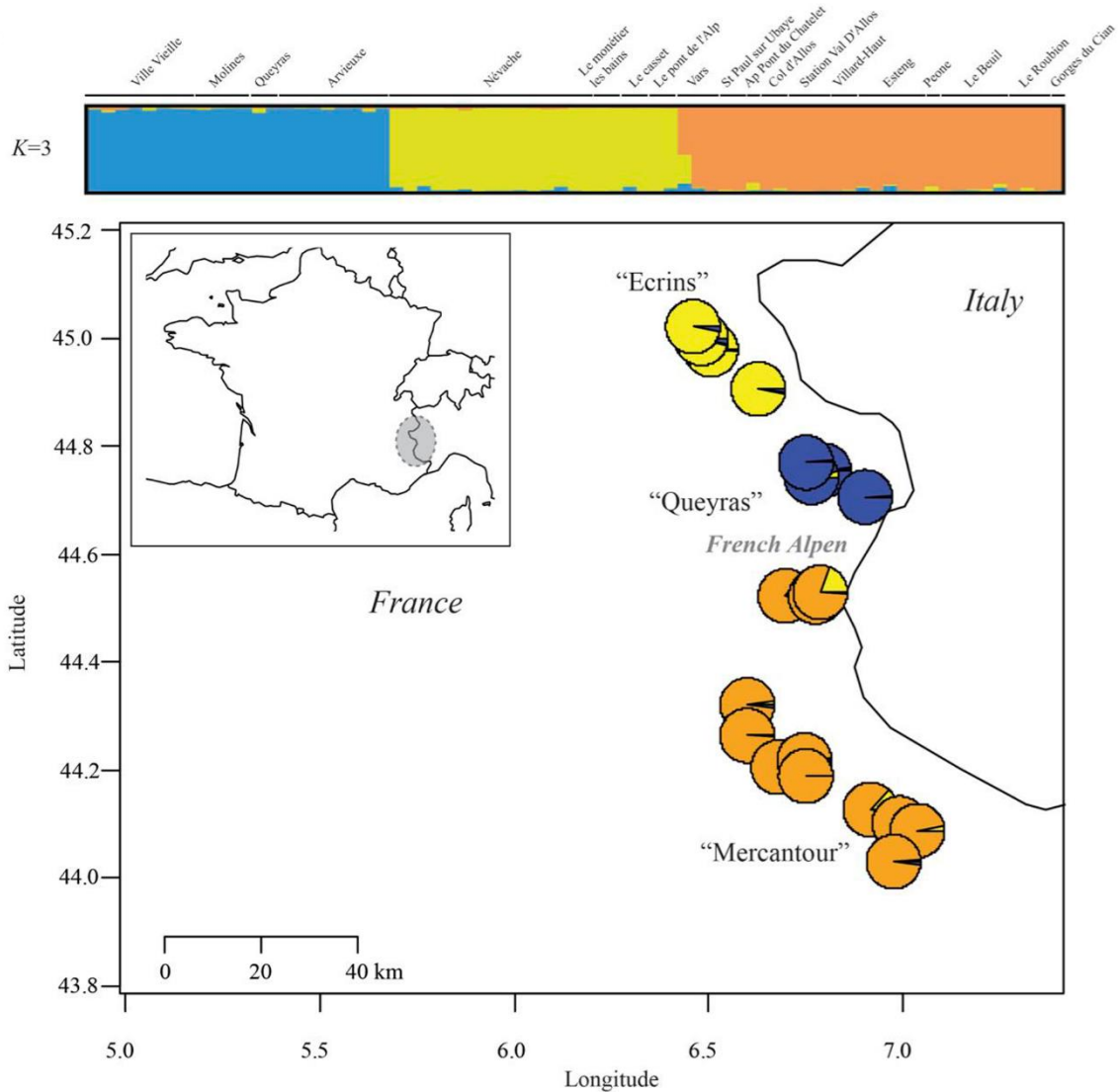
689 **Figure 3. Principal components analysis on *Prunus brigantina*.** Colours refer to the

690 genetic clusters inferred from the STRUCTURE analysis, according to the barplots at *K*=3

691 in Figure S2.

692

693

22

694 **Figure 4. Unrooted weighted neighbour-joining (NJ) tree of *Prunus brigantina* and other**

695 **Prunophora species.** The species are represented by the same colour as the ones used in

696 STRUCTURE barplots (*K*=8, Figure S5). The NJ tree was built with DARwin, bootstrap support

697 values were obtained from 30,000 repetitions. Bootstrap values when greater than 50% are shown

698 above the branches. (*) corresponds to the P0489 plumcot sample. Classification into sections

699 was made according to Krüssmann (1978) and Reales et al (2010).
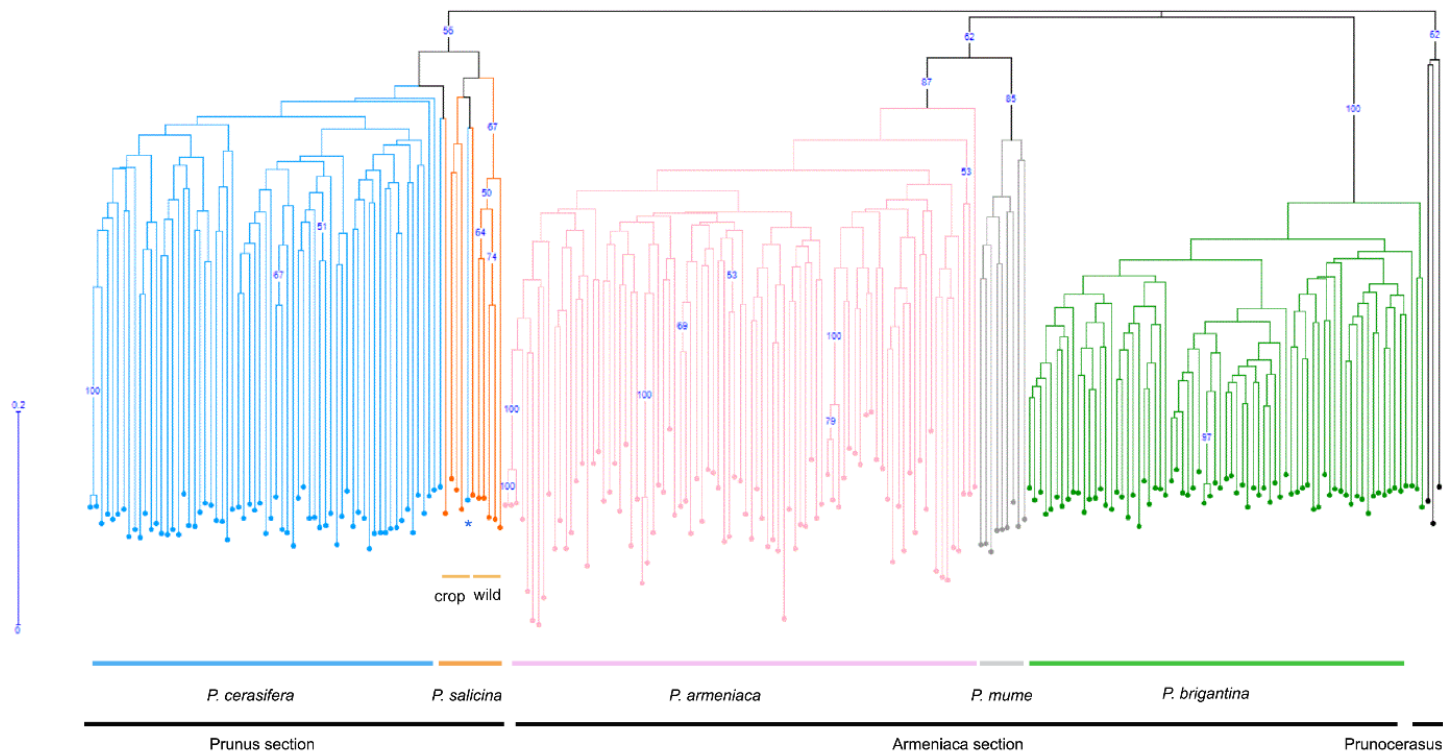
Figure 5. Principal components analysis (PCA) on five Prunophora species performed with DARwin.

The sampling for this analysis included *P. cerasifera* (*N*=66) in blue, *P. armeniaca* (*N*=87) in pink, *P. brigantina* (*N*=73) in green, the Chinese apricot tree *P. mume* (*N*=9) in grey and Japanese plum, *P. salicina* (*N*=10) in orange. Black dots correspond to *Prunocerasus* species (*P. mexicana*, *P. munsoniana* and *P. maritima*). Colours refer to the genetic clusters inferred from the STRUCTURE analysis, according to the barplots at *K*=8 in Figure S5.
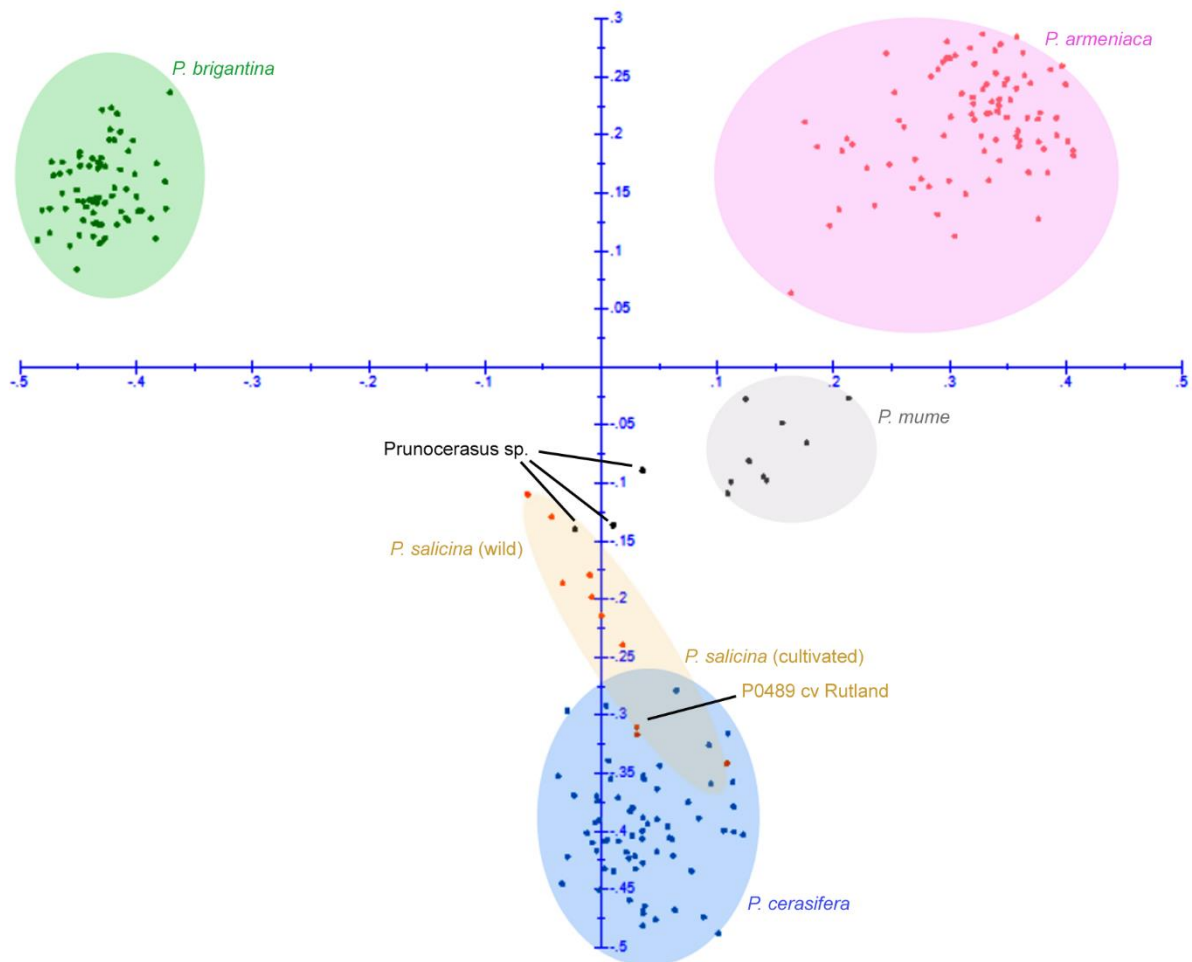
711 **Figure 6. Identification of the core collection of *Prunus brigantina* population based**

712 **on the strategy maximizing allelic diversity.** The genetic diversity in terms of number

713 of alleles (left) or percentage of variation compared to the whole dataset (right) is plotted

714 for different core collection sizes. Details on the accessions retained for each percentage

715 rate are presented in Table S3.

716

717
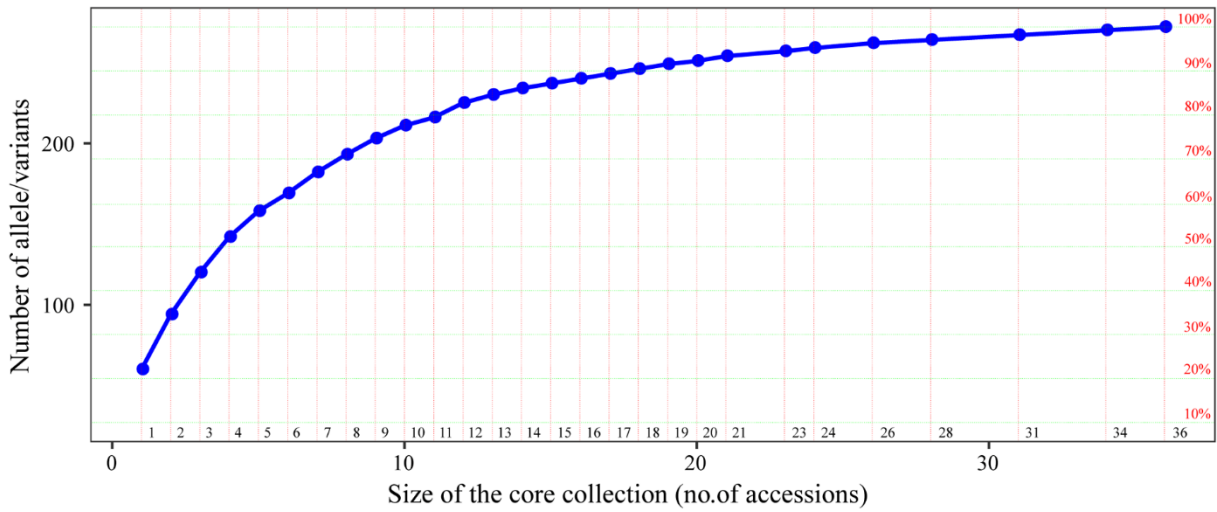
Tables

Table 1. Different datasets including *Prunus brigantina* and other apricot species in this study.

* indicate a *P. brigantina* dataset that includes 71 individuals sampled from the French Alps and 2 samples from the French GRC repository. ^Δ Prunocerasus species are represented by *P. mexicana* (*N*=1), *P. munsoniana* (*N*=1), *P. maritima* (*N*=1), *P. americana* (*N*=1), *P. subcordata* (*N*=1).

Datasets	Description	Number of accessions	Number of accessions for each *Prunus* species under study (*N*)							
			P. brigantina	*P. armeniaca* (wild)	*P. armeniaca* (cultivated)	*P. sibirica*	*P. mume*	*P. mandshurica*	*P. salicina*	*P. cerasifera*
1	*P. brigantina*	73*	73*	-	-	-	-	-	-	-
2	*P. brigantina* and accessions of other Armeniaca species	648	73*	204	270	84	9	8	-	-
3	*P. brigantina* and accessions of the *Armeniaca*, *Prunus* and *Prunocerasus* sections ^Δ	250	73*	-	87	-	9	-	10	66
4	*P. cerasifera*	66	-	-	-	-	-	-	-	66

Table 2. Pairwise population Jost' *D* of *P. brigantina*

Population	Queyras	Ecrins	Mercantour
Queyras	-	0.116	0.14
Ecrins		-	0.097
Mercantour			-